

**Original citation:**

Smith, Oliver, Clapham, Alan J., Rose, Pam, Liu, Yuan, Wang, Jun and Allaby, Robin G.. (2014) Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. Scientific Reports, Volume 4 . Article number 5559. ISSN 2045-2322

**Permanent WRAP url:**

<http://wrap.warwick.ac.uk/62021>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/4.0/>

**A note on versions:**

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: [publications@warwick.ac.uk](mailto:publications@warwick.ac.uk)



<http://wrap.warwick.ac.uk>



## OPEN

SUBJECT AREAS:  
ARCHAEOLOGY  
EPIGENOMICSReceived  
28 March 2014Accepted  
16 June 2014Published  
4 July 2014Correspondence and  
requests for materials  
should be addressed to  
R.G.A. (R.G.Allaby@  
warwick.ac.uk)

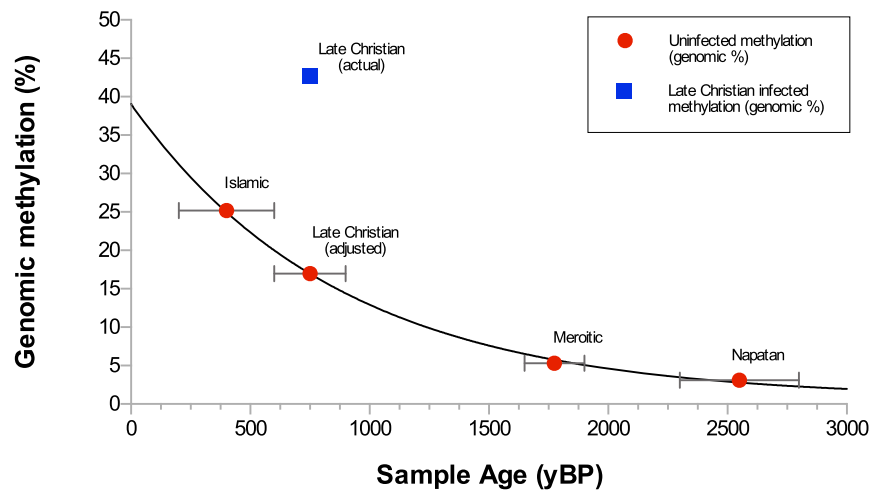
# Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process

Oliver Smith<sup>1</sup>, Alan J. Clapham<sup>1</sup>, Pam Rose<sup>2</sup>, Yuan Liu<sup>3</sup>, Jun Wang<sup>4</sup> & Robin G. Allaby<sup>1</sup><sup>1</sup>School of Life Sciences, Gibbet Hill Campus, University of Warwick, Coventry CV4 7AL, UK, <sup>2</sup>The Austrian Archaeological Institute; Cairo Branch, Zamalek, Sharia Ismail Muhammed, Apt 62/72, Cairo, Egypt, <sup>3</sup>BGI-Europe-UK, 9 Devonshire Square, London, EC2M 4YF, UK, <sup>4</sup>B BGI-Shenzhen, Shenzhen 518083, China.

Genomic methylation is variable under biotic and abiotic stresses in plants. In particular, viral infection is thought to significantly increase genomic methylation with particularly high activity around transposable elements. Here we present the genomic methylation profiles of grains of archaeological barley (*Hordeum vulgare*) from several strata from a site in southern Egypt, from the Napatan to the Islamic periods (800 BCE – 1812 CE). One sample tested positive for viral infection and exhibits an unusually high degree of genomic methylation compared to the rest. A decreasing trend in global methylation levels according to deposition date shows *in-situ* de-methylation of 5-methylcytosine, which can be described as a diagenetic process. This is most likely a deamination mediated de-methylation process and is expected to lead to 5 mC > T base modifications in addition to the C > U modifications due to cytosine deamination, so represents a time-dependent process of DNA diagenesis in ancient DNA.

The detection of methylation signals in ancient DNA offers the opportunity to gain insight into the acclimation of past genomes to environments and to directly examine the evolutionary process of phenotypic plasticity in adapting to environmental conditions<sup>1,2</sup>. The role of dynamic gene network regulation is of particular importance to understanding the evolution of plants under domestication<sup>3,4</sup>. Elevated methylation of genomic DNA in plants is well documented over a variety of biotic and abiotic stresses<sup>5–7</sup>, particularly in response to viral infection<sup>8</sup>. Methylation of genomic DNA, usually by methylation of cytosine residues to form 5-methylcytosine is an epigenetic process that can be inherited transgenerationally in both plants<sup>9</sup> and animals<sup>10,11</sup>. Cytosine methylation in response to stress is carried out by methyltransferase complexes guided by short interfering RNA (siRNA) of ~24 nt<sup>12</sup>. While sequence specificity of siRNAs are not well understood, there is evidence of siRNA sequence conservation between plants<sup>13</sup> and a tendency for retroelement-linked siRNA to be more active during stress responses<sup>14</sup>. Genomic methylation by siRNA activity is thought to inhibit transcription<sup>15</sup> in response to environmental stresses including biotic stress via the RNA-Directed DNA methylation pathway<sup>16,6</sup>.

Cytosine methylation in Pleistocene bone has been detected indirectly<sup>17</sup> and directly through bisulphite sequencing<sup>18</sup>, and led to the suggestion that the methylation signal is very stable over time. More recently, methylation at the genomic level has been quantified in ancient human genomes, as have their *in vivo* effects<sup>19,20</sup> but the rate at which methylated sites become lost over time has not been studied in any detail. The site at Qasr Ibrim offers a good opportunity to study genomic level methylation and diagenetic demethylation in plants, because of the occurrence of multiple archaeological strata in which plant material is present. The site is of further interest due to the remarkable preservation of ancient RNA<sup>21</sup> and DNA<sup>22,23</sup>, compared to the expectations of previous studies into thermal age of bone-originating DNA<sup>24–26</sup>. We believe this condition to be symptomatic of seed endosperm's natural propensity for long-term stasis, and the arid climatic stability at the site<sup>27</sup>. We recently identified the presence of ancient RNA in barley from this site in the presence of an ancient RNA virus genome that likely was an agent of stress at the stratum where it occurred<sup>21</sup>. Furthermore, this stratigraphic series of seeds are all thought to have been under the abiotic stress of low water availability, but locally adapted to these hostile environmental conditions<sup>22</sup>.



**Figure 1 | Total genomic methylation according to sample age.** Negative exponent indicates a methylation half-life of 700.1 years. The observed genomic methylation level of the Late Christian sample (blue point) has been normalised according to previous methylation change data, under the hypothesis that this archaeological sample was infected with Barley Stripe Mosaic Virus (BSMV).

To better understand the stability and decay of methylation in ancient DNA, in this study we aimed to quantify the global methylation profile of archaeological barley over time using samples across multiple strata at the site of Qasr Ibrim, and to use RNA profiling to identify candidate regions for siRNA mediated methylation of loci.

## Results

We found a downward trend of decreasing overall genomic methylation according to the age of the sample, with the single exception of one sample from the Late Christian stratum, Figure 1, Table 1. Interestingly, this was the only sample with genetic evidence of biotic infection from Barley Stripe Mosaic Virus (BSMV)<sup>21</sup>. The level of genomic methylation in this sample was also elevated compared to a modern, uninfected control samples, and likewise the calculated *in vivo* rate of siRNA activity to be highly elevated in the Late Christian sample compared to the control.

**In vivo methylation state of the Late Christian barley.** The Late Christian sample was the only archaeological barley sample in which the presence of BSMV was detected<sup>21</sup>. Viral infection is known to cause increased methylation<sup>8</sup>, so we speculated that this viral infection could be a cause of the high methylation state observed in this sample. Data on genome-wide methylation changes in plants following infection by a ssRNA virus are sparse, so a correction factor was calculated based on the mean genomic methylation increase of the nearest available scenario (2.51 fold in F<sub>1</sub> progeny of *Nicotiana tabacum* following infection with an ssRNA virus)<sup>8</sup>. The observed value of genomic methylation in the Late Christian barley genome (42.6%) was corrected to 16.99%. Although the Qasr Ibrim barley may differ in viral stress response compared to the *Nicotiana tabacum* line due to taxonomy and incomplete inbreeding in antiquity, we were interested to observe that the corrected methylation level is consistent with an exponential probability density function when placed with observed methylation values for three other strata (exponential regression R<sup>2</sup> = 0.9996; Figure 1). This close agreement supports our assertion that the viral infection being the cause of the high methylation signature in the Late Christian is likely to be correct. Like DNA, the half-life of 5 mC is likely to be dependent on tissue type and long-term environmental conditions. Assuming a broadly consistent climatic temperature, we calculated the 5 mC half-life of barley DNA at Qasr Ibrim to be 713.38 years based on methylation levels and stratigraphic dates of the non-infected samples, using equation (1):

$$t_{1/2} = \frac{(t_1 - t_2) \cdot \ln(2)}{\ln\left(\frac{V_{t1}}{V_{t2}}\right)} \quad (1)$$

where  $V_{t1}$  and  $V_{t2}$  are the values of percentage methylation at the Islamic and Napatan strata respectively,  $t_1$  and  $t_2$  are the dates of the respective strata. The estimated half-life was used to infer the methylation level of Late Christian barley at the point of deposition using equation (2):

$$V_{t0} = 2^n V_t \quad (2)$$

where  $V_{t0}$  is the value of percentage methylation at the time of deposition,  $V_t$  is the value of methylation in the Late Christian, and  $n$  is the number of half lives that passed given by:

$$n = (t_0 - t_1) / t_{1/2} \quad (3)$$

The Late Christian sampled was calculated to be 97.78% methylated at the time of deposition, which is consistent with the expected plant response to viral infection<sup>8</sup>. However, we calculate the *in vivo* methylation state of the uninfected Qasr Ibrim barley would have been 38.96%, consistent with the genomic methylation state of 7 modern control samples (mean 35.16%, standard deviation 8.75%; Figure 2). In the control samples we found no obvious correlation between genomic methylation and local environment, although this was not an expectation as there is no supposition of modern cultivars being under stress.

**Demethylation of 5-methylcytosine in ancient DNA.** The reduction in the methylation signal over time that we observe could be

**Table 1 | Sample data including age range of corresponding archaeological period, approximate median age value assigned to each sample, and genomic methylation levels for each sample. Both adjusted and unadjusted values for the Late Christian sample have been included, see text for details**

Sample period	Age range (yBP)	Median age value (yBP)	Genomic methylation (%)
Islamic	200–600	400	25.2
Late Christian	600–900	750	42.64 (16.99 adjusted)
Meroitic	1650–1900	1775	5.3
Napatan	2300–2800	2550	3.12

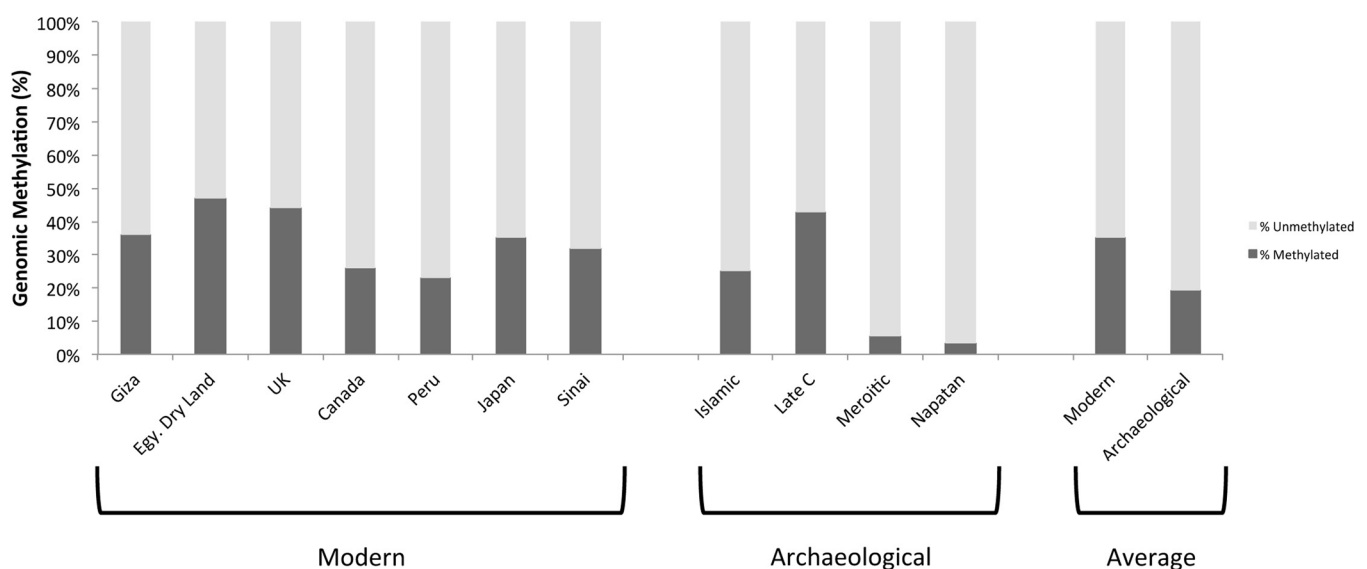


explained by either a chemical modification of the methylated molecule (de-methylation), or by increased fragmentation of DNA molecules over time. In the latter case the smaller DNA molecules captured by beads in the analysis would have fewer methylated cytosine bases because the number of bases on the molecule is less rather than a lower proportion of methylated bases. To investigate this we ascertained through a modeling approach to what extent the observed pattern of DNA degradation could explain the reduction in methylation signal over time. The frequency distributions of DNA fragments from the non-hypermethylated archaeological samples from the Islamic, Meroitic and Napatan strata were inferred from fluorometric measurements using an Agilent Bioanalyzer (Figure S1). The resulting patterns of increased fragmentation over time are consistent with previous patterns observed for desiccated plant remains from of a similar age from this site<sup>23</sup>. We simulated the expected range of methylated cytosines to be sampled from a 1000 molecules from each of the archaeological samples (Figure S2a), and the consequent expected range of reduction in the methylated cytosines sampled between archaeological strata (Figure S2b). We found that fragmentation could account for an 80% of reduction in methylation signal between the Islamic and the Meroitic strata, and 70% of the reduction both between the Islamic and Napatan strata and the Meroitic and Napatan strata. These data support the notion that the contribution of DNA fragmentation to the loss of methylation signal stabilizes to a value of about 70% over time. This analysis assumes that both large and small DNA molecules would be equally likely to bind to beads during the methylation assay. The methyl-binding domain (MBD) bead-based method we used does not appear to discriminate in favour of large fragments during enrichment for target methylated molecules unlike column-based assays. In fact it has been suggested that smaller fragments are more efficiently captured using this method<sup>28</sup>, suggesting that our estimates of the contribution of DNA fragmentation to the reduction in methylation signal are too high. Therefore in all cases examined the extent of DNA fragmentation was insufficient to explain all the reduction in methylation signal. We therefore conclude that a chemical process resulting in apparent de-methylation significantly contributes to the loss of methylation signal over time.

Several mechanisms of cytosine de-methylation have been proposed, including spontaneous deamination<sup>29</sup>, hydrolysis<sup>30</sup> and *in vitro* processes largely involving base excision repair machinery<sup>31</sup>. Since *in vivo* processes in archaeological material can be dismissed,

spontaneous or direct chemical processes are most likely. The spontaneous loss of the 5' methyl group from 5-methylcytosine is not well understood or chemically defined<sup>32</sup>. The more likely spontaneous event is deamination of 5-methylcytosine to thymine in the presence of water, and has been documented as a process previously<sup>33</sup>. While not strictly a de-methylation event since the 5' methyl group remains intact in the resultant thymine, 5 mC > T reactions would present as non-methylated sites following sequencing. 5 mC deamination reactions have a 2–4 fold higher rate than cytosine > uracil reactions<sup>30</sup> and so are likely to occur at greater frequency in genomes with higher methylation. This study therefore provides empirical evidence that observed C > U/T transitions in ancient DNA are the result of two separate processes of deamination, the second of which is specific to methylated cytosines as has previously been suggested<sup>17,19,20</sup>. Samples that are more methylated, for instance for reasons of stress, may therefore have a stronger C to T bias. Verification of this possibility warrants further investigation to correlate methylation profiles with age and C to T/U bias profiles from DNA sequence data.

**Methylating siRNA in archaeological barley.** We hypothesized that methylation-mediating siRNA would occur at greater frequency in a plant genome undergoing stress. Candidate regions of siRNA interference in the genome are expected to be associated with localized points of high levels of read depth, which should differ from the general background of depth produced by mRNA breakdown templates. We identified the top 6 loci (AGL97, AP2L1, Lr34, Mla, eIF4E and Sukkula-retroelement-containing locus) that contained the largest peaks of read coverage in our data set generated from the Late Christian archaeological sample, Figure S3. Of particular interest in this candidate list is the eIF4E locus that has previously been shown to be involved in pathogen response<sup>34</sup>, given that we previously showed the Late Christian sample to have been under pathogen attack<sup>21</sup>. We compared the coverage depth of reads at this locus in the four modern samples with that of the archaeological Late Christian sample, figure S4. It is problematic to compare peak height relative to background between modern and archaeological samples to interpret which peak represents the greater potential siRNA activity. This is because the extent to which there is a greater background of mRNA degradation products in the archaeological sample relative to the modern samples is difficult to accurately determine. We therefore compared coverage peaks at the eILF4 locus between the modern and archaeological samples simply



**Figure 2 | Total genomic methylation of modern and archaeological samples.** No correlation was identified between geographic region, climate and methylation level. The projected methylation state of ‘modern’ Qasr Ibrim barley is similar to observed levels of existing barley.





for presence or absence of peaks. The modern samples all showed numerous peaks within a narrow range of magnitude, while the archaeological sample showed two peaks that dominated the profile (Figure S4). We interpreted this result as evidence supporting the notion of an increased siRNA action in the archaeological sample, and therefore a good candidate region for siRNA mediated methylation.

We investigated whether the coverage peaks at the eIF4 locus in the archaeological barley were associated with increased methylation by performing a methylation-specific (bisulfite) PCR along regions of the genome showing relatively high coverage in the archaeological sample. We selected 'peak 1' (Figure S3, S4) where coverage in the archaeological sample accounted for 10% of the total reads assigned to that locus, and where equivalent coverage in the control samples was completely absent. Primers were designed to cover a 105 bp amplicon along this locus. After bisulfite treatment, the archaeological sample exhibited no C>U modification to cytosine or the antisense strand equivalent (G>A) indicating complete methylation of the amplified locus (Figure S5). Conversely, the same locus in the Sinai control barley showed partial G>A transitions at two nucleotide sites indicating only partial *in vivo* methylation. A secondary control on an archaeological sample from a different stratum (Qasr Ibrim Islamic) showed complete demethylation at three sites and partial methylation at one, further supporting the Late Christian sample's elevated genomic methylation *in vivo*.

## Discussion

These data lead us to conclude that de-methylation of cytosines in ancient DNA exists as a diagenetic process. We conclude that the half-life value calculated for the loss of signal in the methylation assay is a result of both the DNA fragmentation process and the diagenetic de-methylation of cytosines over time, the rates of which are likely dependent on environmental conditions and tissue type. The proportion of the cytosines that have become de-methylated ( $C_{dm}$ ) on the remaining fragment sizes can be calculated by taking account of the proportion of the signal loss due to fragmentation in equation (4):

$$C_{dm} = \frac{1-p}{2-p} \quad (4)$$

Where  $p$  is the proportion of the signal that is lost due to DNA fragmentation. The experimental data suggest an over all half life of 713.38 years for the loss of methylation signal, and our simulations indicate that about 70% of the signal loss is attributable to DNA fragmentation. Therefore over the course of one observed half-life of reduced methylation signal we expect that 23% of the cytosine residues on the remaining fragments will have become de-methylated through a chemical attack mechanism, which is 0.46 of a half-life of the over all signal loss. Therefore we estimate the half-life of the cytosine de-methylation process to be in the order of 1522 years in these samples, although may be significantly longer for similar samples in colder environments as observed<sup>19</sup>. We hypothesize that the chemical mechanism of cytosine de-methylation would lead to a C > T base modification as with the normal cytosine deamination associated with ancient DNA and so represents a new contributing factor to base modification. We predict that a highly methylated state would be expected to contribute an elevation in the C > T/U base modification rate. In the case of the Late Christian sample, assuming a 97% level of genomic methylation at deposition, in every 100 bases of ancient DNA that had methylated to that level, six C > T transitions would be attributable to de-methylation. A sample of the same age and preservation that had a 40% level of methylation at deposition would be expected to have two C > T transitions over 100 bases. This half-life value of cytosine de-methylation contrasts with the high level of stability previously reported for Pleistocene bone<sup>17,18</sup>, which is likely a reflection of the significantly higher temperature at Qasr Ibrim compared to permafrost conditions and consequent greater

thermal age. While temperature is the overriding determinant of DNA preservation over time<sup>24–26</sup>, the further environmental factors of low humidity and constancy of climate at this site are also likely to have influenced the half-life. These secondary factors are likely to have a greater influence on RNA than DNA preservation, which may limit the preservation contexts in which RNA seq and methylation data may be combined. Further studies of the methylation half life in a wider variety of contexts are required to better understand this diagenetic process.

We also conclude that the underlying causes of *in vivo* methylation can still be seen from inferring siRNA action from aRNA profiles and their corresponding DNA methylation patterns, given the right preservation conditions. Furthermore, we conclude that biologically relevant information can be extrapolated from ancient RNA data, allowing a new dimension in the study of ancient biomolecules. These findings suggest new potential in archaeogenomics to complement research into what an archaeogenome *is*, with what an archaeogenome *does*.

## Methods

**For methylation analysis**, archaeological barley seeds from Qasr Ibrim were selected from four strata: Islamic (accession 84/102, 1400–1812 CE), Late Christian (accession 84/71, 1100–1400 CE), Meroitic (accession 84/46, 100–350 CE) and Napatan (accession 84/165B, 800–300 BCE). The same DNA extraction and methylation analysis methods were applied to all four samples. Control samples for methylation analysis were obtained from the USDA Germplasm bank on the basis of climatic distribution: northern Egypt (x2; Giza PI 6426786 and Dryland PI 542707), UK (PI 531846), Canada (Clho 6048), Peru (PI 510559), Japan (PI 1862613) and the Sinai peninsula (PI 564601). RNA for Illumina sequencing was extracted from three seeds of the Late Christian stratum and one seed from four of the control samples (Japan, Egyptian Dryland, Peru and Sinai).

**DNA extraction.** DNA was extracted from archaeological material using a modified CTAB extraction protocol. Seeds were crushed to powder, immersed in 2% CTAB buffer (containing 1% PVP) and incubated at 37°C for 5 days. DNA was extracted once with chloroform:isoamyl-alcohol 24:1 and purified using the DNEasy Plant Mini Kit (Qiagen) with a single modification to binding buffer amount (3x eluant volume as opposed to suggested 1.5x). DNA was extracted from modern material by crushing, immersion in 2% CTAB buffer (containing 1% PVP) and incubation for 1 hour at 65°C. DNA was extracted once with chloroform:isoamyl-alcohol 24:1 and purified using the DNEasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions.

**Methylation assay.** DNA from modern samples was sheared by sonication into ~300 bp fragments using a Bandelin Sonopuls UW2070 (5 cycles of 30 seconds sonication at 30% followed by 30 seconds cooling on ice). DNA from archaeological samples were not sheared due to the fragmented nature of ancient DNA. Enrichment of methylated DNA was performed using the Methylminers Methylated DNA Enrichment Kit (Applied Biosciences), according to the manufacturer's instructions. DNA concentration (and calculated total amount) for each sample was quantified using the Qubit platform (Invitrogen) before and after enrichment.

**RNA extraction.** RNA from archaeological samples was extracted using a modified protocol of the MirVana miRNA Isolation Kit (Ambion). Seeds were crushed to powder, immersed in 2% CTAB buffer (containing 1% PVP) and incubated at 37°C for 5 days. Total nucleic acids were extracted once with chloroform:isoamyl-alcohol 24:1 and precipitated with 3 volumes ethanol in 0.3 M NaOAc. The resulting pellet was washed twice with 80% ethanol and dissolved in guanidium salt-based binding buffer. RNA was isolated by extraction with phenol-chloroform pH 4.2 and subjected to on-column purification. RNA fragments of <300 nt were enriched according to the manufacturer's instructions. RNA from modern samples was extracted similarly, with a reduced initial incubation phase (65°C for 1 hour). Pellets from modern samples were dissolved in SSTE buffer to remove excess polysaccharides and re-precipitated before dissolution in guanidium salt buffer. RNA fragments were enriched for small fragments <300 nt.

**Illumina libraries.** Illumina small RNA libraries for the Late Christian (archaeological) and Sinai (modern) samples were built and purified using an Air Small RNA Library Kit (Bioo Scientific) according to manufacturer's instructions. Sequencing took place on the HiSeq 2500 platform. Libraries for additional control samples were built and purified using a NextFlex Small RNA Sequencing Kit (Bioo Scientific) to allow multiplexing and sequencing on the MiSeq platform. Sequences were deposited in the read archives of NCBI BioProject SUB445675.

**Bisulfite treatment.** DNA from archaeological Late Christian, and Islamic strata and the modern Sinai sample was treated using the Methylcode Bisulfite Conversion Kit (Invitrogen) according to manufacturer's instructions. PCR primers were designed to



give amplicons ~100 bp to compensate for the degenerative effects of strand breakage associated with ancient DNA and bisulfite treatment.

**Bioinformatics.** Total small RNA datasets were aligned to the barley genome using Bowtie 0.28. Loci with a coverage depth >100 at any point were assigned for further analysis. Individual reads aligning to each locus in bowtie were exported in Fasta format using a shell script of the author's design. Genbank files of corresponding loci were loaded into Geneious R6 (Biomatters), and exported reads aligning to each locus were re-aligned using the Geneious align feature. Coverage data was exported in ACE (PHRAP) format and imported into Tablet. Corresponding data was exported from Tablet into readable format and parsed to give position/coverage values using a shell script of the author's design. Position and coverage value information was charted using Excel. Locus detail from Genbank format was exported from Geneious as a PDF and overlaid onto the corresponding Excel chart. Sequence traces pre- and post-bisulfite treatment were inspected visually to give three 'grades' of methylation inference (zero, partial and complete).

The exponential decay curve of demethylation in archaeological materials was calculated by assigning each sample an x-value of approximate age, based on the median value of archaeological period (after<sup>23,27</sup>, Table 1) to the nearest 25 years, and a y-value of genomic methylation. The Late Christian sample was adjusted to its hypothesized uninfected level as described earlier. An  $R^2$  value and plot were then generated by running a nonlinear regression analysis with one-phase decay of the data, using Graphpad Prism 6.0c for Mac OSX. A second series of the actual methylation level of the Late Christian sample was also plotted for comparison. The half-life of genomic methylation for archaeological samples in this context was then calculated as described earlier.

The distribution of fragment sizes in DNA extracts from archaeological samples was inferred from Bioanalyzer output data. The Bioanalyzer output gave values of absorbance associated with DNA over time as the sample moved through a capillary. Markers of 35 bp and 10,800 bp were used to calibrate the time associated with intervening DNA sizes using a semilog interpolation in GraphPad Prism 6.0c for Mac OSX, which produced a best fit curve of size (y) against time (x). Values of time from the original Bioanalyzer output were then converted to size in base pairs using the conversion formula:

$$x_i = 10^{((t_i - y)/m)} \quad (5)$$

where  $x_i$  is the number of inferred base pairs at time point  $t_i$ ,  $y$  is the y axis intercept and  $m$  is the gradient of the semilog interpolation of size against time.

The resulting frequency distributions of DNA size for each sample were used to construct a non-parametric probability density as input for a simulation to determine the effect of DNA fragmentation on the cytosine methylation signal. In the simulation 1000 molecules were randomly sampled from the probability density. The total number of methylated cytosine bases were then determined assuming methylated cytosines occurred at a frequency of 10%, representing a 40% methylated genome as an approximation of a 'modern' Qasr Ibrim barley. Each simulation was repeated 1000 times, giving a frequency range of expected methylated cytosines. A frequency range of the expected loss of signal due to DNA fragmentation was then calculated by dividing each of the simulation outputs from one distribution by each of the simulation outputs of the second distribution representing a complete sampling of combinations of numbers of methylated cytosines sampled. All simulations and subsequent dividing of output frequency distributions was carried out using Perl scripts of the last author's design.

- Price, T. D. & Qvarnström, I. D. E. The role of phenotypic plasticity in driving genetic evolution. *Proc. Roy. Soc. Ser. B* **270**, 143–144 (2003).
- West-Eberhard, M. J. Evolution in the light of developmental and cell biology, and vice versa. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8417–8419 (1998).
- Piperno, D., Holst, I., Winter, K. & McMillan, O. Teosinte before domestication: Experimental study of growth and phenotypic variability in Late Pleistocene and early Holocene environments. *Quaternary International* DOI:10.1016/j.quaint.2013.12.049 (2014).
- Allaby, R. G. *et al.* Using archaeogenomics and computational approaches to unravel the history of local adaptation in crops. *Phil. Trans. Roy. Soc. B*. in press (2014).
- Boyko, A. *et al.* Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. *PLoS ONE* **5**, e9514 (2010).
- Dowen, R. H. *et al.* Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. USA* **109**, E2183 (2012).
- Verhoeven, K. J. F., Jansen, J. J., van Dijk, P. J. & Biere, A. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* **185**, 1108–1118 (2010).
- Boyko, A. *et al.* Transgenerational changes in the genome stability and methylation in pathogen-infected plants. *Nucleic Acids Res.* **35**, 1714–1725 (2007).
- Paszkowski, J. & Grossniklaus, U. Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Curr. Opin. Plant Biol.* **14**, 195–203 (2011).
- Dolinoy, D. C., Huang, D. & Jirtle, R. J. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13056–13061 (2007).

- Heijmans, B. T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17046–17049 (2008).
- Khraivesh, B., Zhu, J.-K. & Zhu, J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *BBA - Gene Reg. Mech.* **1819**, 137–148 (2012).
- Heisel, S. E. *et al.* Characterization of unique small RNA populations from rice grain. *PLoS ONE* **3**, e2871 (2008).
- Ito, H. *et al.* An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **472**, 115–9 (2011).
- Boyes, J. & Bird, A. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**, 1123–1134 (1991).
- Wassenegger, M., Heimes, S., Riedel, L. & Sanger, H. L. (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76**, 567–576 (1994).
- Briggs, A. W. *et al.* Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucl. Acids Res.* **38**, e87 (2010).
- Llamos, B. *et al.* High-resolution analysis of cytosine methylation in ancient DNA. *PLoS ONE* **7**, e30226 (2012).
- Pedersen, J. S. *et al.* Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Res.* **24**, 454–466 (2013).
- Gokhman, D. *et al.* Reconstructing the DNA methylation maps of the Neanderthal and the Denisovan. *Science* **344**, 523–527 (2014).
- Smith, O. *et al.* A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. *Sci. Rep.* **4**, DOI:10.1038/srep04003 (2014).
- Palmer, S. A., Moore, J. D., Clapham, A. J., Rose, P. & Allaby, R. G. Archaeogenetic evidence of ancient Nubian barley evolution from six to two-row indicates local adaptation. *Plos One* **4**, e6301 (2009).
- Palmer, S. A. *et al.* Archaeogenomic Evidence of Punctuated Genome Evolution in Gossypium. *Mol. Biol. Evol.* **29**, 2031–2038 (2012).
- Allentoft, M. E. *et al.* The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *P. Roy. Soc. Lond. B. Bio.* **279**, 4724–4733 (2012).
- Schwarz, C. R. *et al.* New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Res.* **37**, 3215–3229 (2009).
- Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C. & Collins, M. J. The thermal history of human fossils and the likelihood of successful DNA amplification. *J. Hum. Evol.* **45**, 203–217 (2003).
- Clapham, A. J. & Rowley-Conwy, P. A. New discoveries at Qasr Ibrim, Lower Nubia. In Cappers, R.T.J. ed, *Fields of Change: Progress in African Archaeobotany*. Barkhuis, Groningen, pp 71–77 (2007).
- Aberg, K. A. *et al.* High quality methylome-wide investigations through next-generation sequencing of DNA from a single archived dry blood spot. *Epigenetics* **8**, 542–547 (2013).
- Ehrlich, M., Zhang, X.-Y. & Inamdar, N. M. Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues. *Mutat Res/Reviews in Genetic Toxicology* **238**, 277–286 (1990).
- Shen, J.-C., Rideout, W. M. & Jones, P. A. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucl. Acids. Res.* **22**, 972–976 (1994).
- Zhu, J.-K. Active DNA Demethylation Mediated by DNA Glycosylases. *Ann. Rev. Genet.* **43**, 143–166 (2009).
- Patra, S., Patra, A., Rizzi, F., Ghosh, T. & Bettuzzi, S. (2008) Demethylation of (Cytosine-5-C-methyl) DNA and regulation of transcription in the epigenetic pathways of cancer development. *Cancer Metastasis Rev.* **27**, 315–334 (2008).
- Labet, V., Morell, C., Cadet, J., Eriksson, L. A. & Grand, A. Hydrolytic Deamination of 5-Methylcytosine in Protic Medium: A Theoretical Study. *J. Phys. Chem. A* **113**, 2524–2533 (2009).
- Charron, C. *et al.* Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. *Plant J.* **54**, 56–68 (2008).

## Acknowledgments

We would like to thank the Egyptian Exploration Society for access to the material from Qasr Ibrim. This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC DTA BBG0177941).

## Author contributions

O.S. carried out the research, R.G.A. designed the research project, A.C. and P.R. provided archaeological materials, Y.L. and J.W. provided sequencing facilities, O.S., R.G.A., A.C. and P.R. analyzed the data and O.S. and R.G.A. wrote the manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Smith, O. *et al.* Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. *Sci. Rep.* **4**, 5559; DOI:10.1038/srep05559 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if

the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>